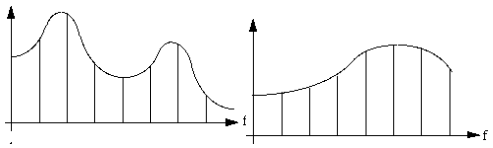
Speech Processing 2011/12

1st Test

March 19th 2012

Please identify this form with your name and student number in the reserved space at the bottom. The answers to multiple-choice questions will only be accepted if inserted in the appropriate place. Wrong answers will be penalized. The phonetic symbols should use the SAMPA alphabet (Lisbon accent).

1. Classify as True (T) or False (F)
   (a) Critical bandwidths in the Mel scale decrease as the center frequencies of the filters increase
   (b) The cepstrum of a pulse train with pulses every N samples is only non-zero at integer multiples of N.
   (c) Consider a pole-zero filter of impulse response $h[n]$, whose roots are all outside the unit circle. The corresponding cepstrum is non-zero only for negative values of n.
   (d) From a computational point of view, the autocorrelation method of linear prediction is less efficient than the covariance method.
   (e) Low values of the LPC order cause the filter to attempt to model individual harmonics.
   (f) The upper part of the vocal folds is more flexible and opens and closes before the lower part.
   (g) Pitch detection algorithms perform better in voiced with voices with high jitter.
2. Classify the initial phoneme of the English word *theory* in terms of
   (a) manner of articulation
   (b) place of articulation (labial, dental, alveolar, velar, platal)
3. Name a phonation type with the highest F0, and another one in which a part of the glottis remains open.
4. Name the three most important prosodic components.
5. Neutral vowels in all languages are front, center or back?
6. Figure 1 shows two diagrams. The one on the left is a schematic picture of the spectrum for a particular configuration of the vocal tract. To simplify, let us assume that it is the neutral vowel. The one on the right corresponds to the same configuration, but the vocal tract is filled with helium, instead of air. The second resonance has been shifted right off scale. What changes in the equation for the formant frequencies to justify this shift?



1a 1b

7. Consider the following sentence:
   As farmácias são obrigadas a ter sempre disponíveis os medicamentos mais baratos.
   (a) Write the broad phonetic transcription.
   (b) Which of the following classes is less represented in this sentence?

- vowels
- fricatives
- plosives
- liquids

8. Identify the 10-digit sequence by inspecting the spectrograms and waveforms in pages 3 and 4. There are no repeated digits. The recordings correspond to telephone speech.

9. Compute how many LSP coefficients are computed in total, during the analysis of a sentence of duration 2s. The LSP analysis windows are 20 ms long, and overlap by 50%. The order of the LPC analysis is 12.

10. What is the zero crossing rate of a sinusoidal signal of F0 frequency, sampled at Fs.

11. Figure 2a presents the short-time Fourier transform (magnitude) of a vowel segment, including its LPC spectral envelope.

    (a) Indicate approximate values for F0, F1 and F2 (Hz).

    (b) To which of the 3 vowels of the extremes of the vowel triangle does it correspond?

    (c) May it correspond to a window of 5 ms?

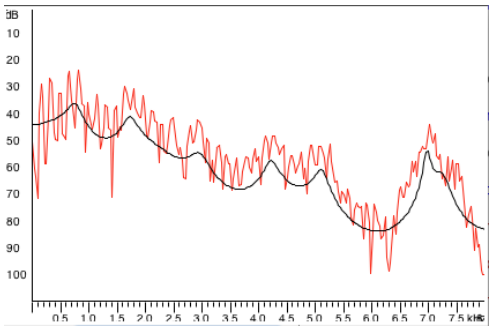
    (d) May it correspond to a deep male voice?

    (e) May the spectral envelope be obtained with an LPC analysis of order 8?

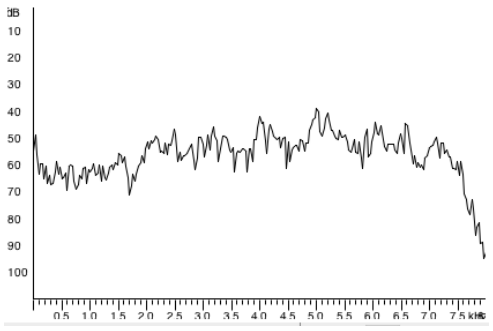
    (f) Does it correspond to narrowband speech?

    (g) Figure 2b presents another short-time magnitude spectrum obtained with the same utterance. May it correspond to the residual of a vowel segment?

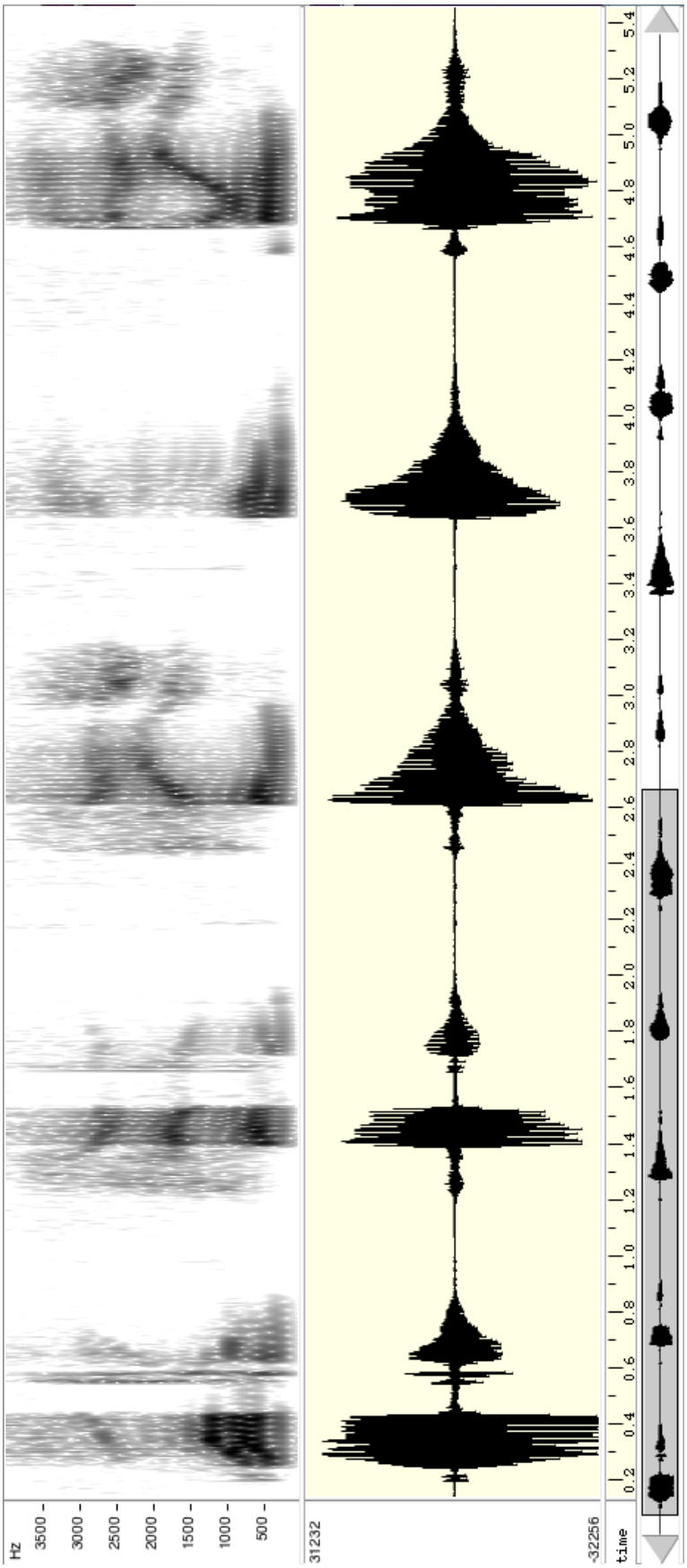
    (h) May it correspond to a fricative sound?

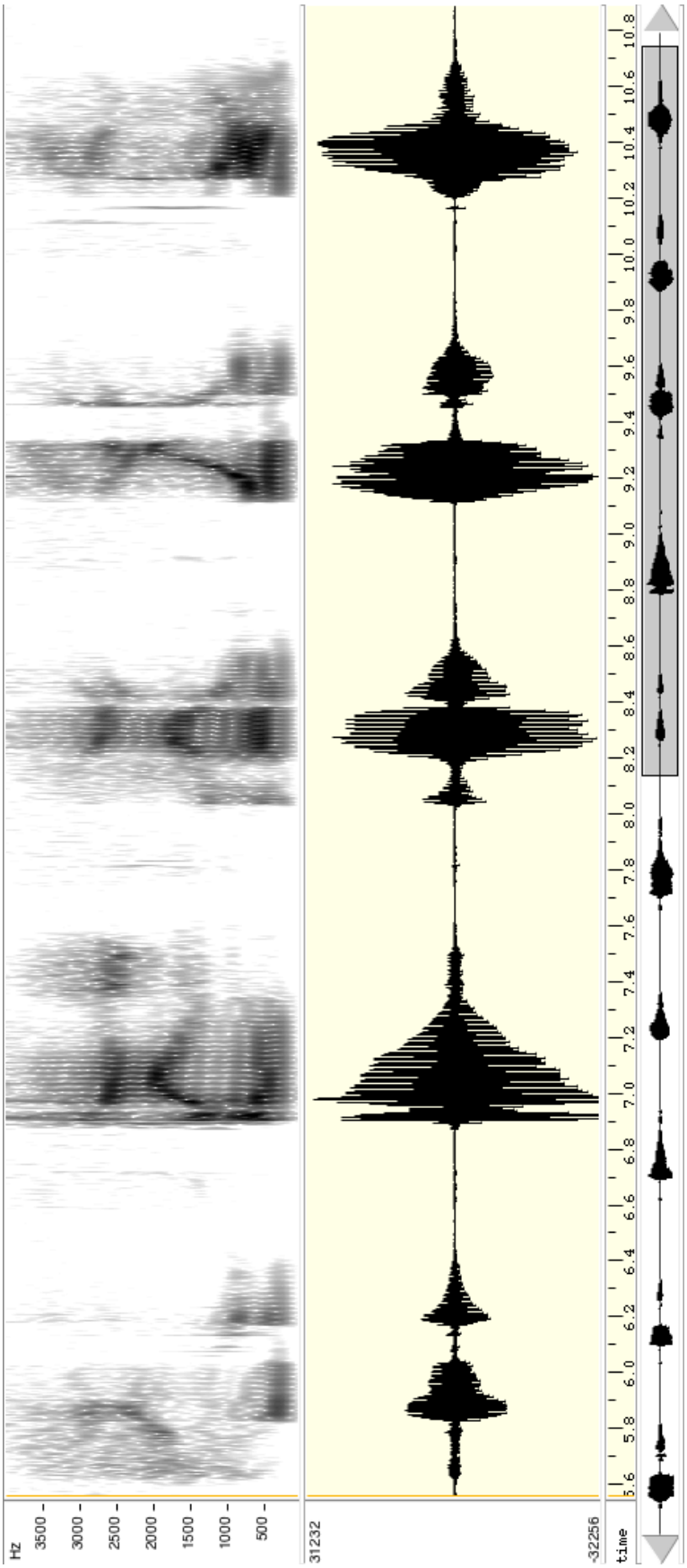


8a



8b


Hz
3500
3000
2500
2000
1500
1000
500
31232
-32256
time
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8
3.0
3.2
3.4
3.6
3.8
4.0
4.2
4.4
4.6
4.8
5.0
5.2
5.4


Hz
3500
3000
2500
2000
1500
1000
500
31232
-32256
time
5.6
5.8
6.0
6.2
6.4
6.6
6.8
7.0
7.2
7.4
7.6
7.8
8.0
8.2
8.4
8.6
8.8
9.0
9.2
9.4
9.6
9.8
10.0
10.2
10.4
10.6
10.8

Test 1 - Answers

Name:	
Number:	

1. (1.4 val.) Indicate T or F:

a	b	c	d	e	f	g

2 (1.2 val.)

manner	
place	

3 (1.6 val.)

a	
b	

4 to 6. (1.5/0.8/2.0)

4	
5	
6	

7. (2.0 val.)

(a)	
(b)	

8. (3 val.)

9 to 10. (2.0/1.0)

9	
10	

11. (3.5 val.) Complete:

F0	F1	F2	Vow	c	d	e	f	g	h